Dear Sir:

Recently Rappaport and Kromhout [AIHAJ 54:654 (1993)] explored the "between worker" exposure variability of 183 homogeneous exposure groups (HEGs) with serial personal air sampling, including more than 15,000 measurements. They showed that about 80% of the groups exceeded an ad hoc upper limit of at most two-fold difference among 95% of the individual mean exposures (${}_BR_{95} = 2$, suggested by Rappaport 1991 page 101) and concluded that most HEGs in industrial hygiene are not homogeneous at all.

Rappaport's conclusions are based on a belief in the general validity of the ${}_BR_{95} = 2$ upper limit for the "between worker" variance in HEGs. As will be demonstrated, ${}_BR_{95} = 2$ cannot be generally valid, because its value is dependent on the number of workers per HEG and the number of measurements per worker. An analogy with dicing will help to get this picture clear.

Further, if the well-established techniques for (1) forming HEGs, (2) examining the lognormal shape of measurements within an HEG, and (3) testing compliance with the limit are used correctly (Hawkins 1991), then the analysis of between-worker variance seems superfluous.

Workers in an HEG and dicers both experience, on the population level, equal probabilities of receiving specific values (concentration or a total number of eyes, respectively). In a limited series, however, values of individuals will differ by chance. I have quantified, in two examples, the differences using a Monte Carlo simulation. In the first example k = 5 dicers throw the die N = 2 times each. The expected total number of eyes $\bar{x}_{E,i}$ (i = 1,2,...5) in an average game is in ascending order 4, 6, 7, 8, 10, respectively (established with 1000 series from a random generator and rounded to the nearest integer). The extreme values $\bar{x}_{E,1}$ and $\bar{x}_{E,k}$ differ by more than a factor of two on the average, and this is completely due to chance!

In the second example the number of dicers is expanded to k = 10 and the number of throws to N = 5. This results in $\bar{x}_{E,i}$ = 12, 14, 15, 16, 17, 18, 19, 20, 21, 23 (again rounded to the nearest integer). If the number of throws per dicer expands to infinite, then the average number of eyes will, for all dicers, asymptotically tend to the population mean $\mu = \bar{x}_E \approx 3.5$ per throw (Central Limit Theorem). So the difference between the number of eyes per dicer is inherent to limited serial dicing (and on this principle many games using dice are based).

Since the distribution of exposure in an HEG is best represented by the lognormal model, the geometric mean (GM) is most often used as the measure of location. The GMs of k workers within an HEG will differ if the sample size per worker is limited. GMs will asymptotically tend to the unknown population geometric mean EXP(μ) if the size increases to infinite (Central Limit Theorem).

In the following two examples I will quantify the at-random difference (= nonsystematic inhomogeneity) between samples with serial measurements, which are drawn from the same population base. The procedure is equal to that used in the dicing experiment. Random standard normal deviates are generated using the direct method from Abramowitz (1970 page 953), using EXP(μ) = 1 ppm as the location and EXP(σ) = 2.71828 as the scale parameter.

In the first example k = 5 workers are sampled N = 2 times each. The expected geometric means $GM_{E,i}$ (i = 1,2,...5) in the average sampling plan (calculated from 1000 series) are in ascending order 0.4, 0.7, 1.0, 1.4, and 2.3 PPM, respectively. The extreme values GM_1 and GM_5 differ by more than a factor of five on the average, and this is completely due to chance!

The average "within worker" variance of the 5000 pairs is about ${}_wGSD \approx 2.3$ (GSD = geometric standard deviation). Because the total variance was fixed at EXP(σ) = 2.71828, the nonsystematic "between worker" variance can be estimated as ${}_BGSD_E = 1.74$ and ${}_BR_{95,E} = 8.75$ (using $\sigma_B^2 = \sigma^2 - \sigma_W^2$). Although the differences are completely random, this situation should, according to Rappaport, be rejected as nonhomogeneous!

To quantify the influence of expanding the number of workers and the samples per worker, a scenario was chosen that is comparable with the median situation found in the Rappaport database, namely k = 10 workers with N = 5 measurements each. The expected geometric means $GM_{E,i}$ (i = 1,2,...10) in the average sampling plan (also calculated from 1000 series) are in ascending order 0.5, 0.64, 0.75, 0.85, 0.95, 1.06, 1.18, 1.33, 1.55, and 1.95 ppm, respectively. On average the GM ratio between the extremes in a situation with k = 10 elements and N = 5 samples per element decreases to $GM_{10}/GM_1 = 4$.

Because the average "within worker" variance of the 10,000 quintets is about ${}_wGSD \approx 2.5$, the nonsystematic "between worker" variance decreases to ${}_BGSD_E = 1.49$ and ${}_BR_{95,E} = 4.77$. This value is quite comparable with the observed median situation displayed in Table 4 of Kromhout (1993) for the chemical gases-vapors and in Figure 2 of Rappaport (1993) for the overall situation.

Based on Monte Carlo simulations for every combination "k workers, N measurements, EXP$(\sigma)_T$" an expected ${}_BGSD_E$ can be established by Monte Carlo simulation. In general, ${}_BGSD_E$ will asymptotically tend to zero if k and N increase to infinite. An upper limit of ${}_BR_{95} = 2$ will be a very special case for a situation with high values for k and N.

Statisticians should be able to construct a k,N-dependent confidence interval for ${}_BGSD_E$, to which the observed ${}_BGSD_O$ can be tested (see Land 1988 para. 4.2, page 99). Since the median number of workers, k, and the median number of measurements per HEG, N, in the database were somewhat smaller than those used in the second Monte Carlo experiment, it would not surprise me, if most of the 183 ${}_BGSD_O$ in the Rappaport database would be situated within the confidence limits of ${}_BGSD_E$.

Other possibilities to test inhomogeneity of multiple serial sampling in HEGs are (1) using the $GM_{O,i}$ and $GM_{E,i}$ values in a probability plot (Hawkins 1991), or (2) using the chi-square type test (provided that $\bar{x}_{O,i}$ is normally distributed with equal $GSD_{O,i}$:

$$X^2_{df=k-1} = \sum_{i=1}^{i=k} \frac{\left(\frac{\bar{x}_{O,i}}{\bar{s}_O} - \bar{x}_{E,i}\right)^2}{\bar{x}_{E,i}}$$

with $\bar{x}_i = LOG(GM_i)$ and $\bar{s}_O = \log(\overline{GSD}_O)$

Extreme low values of χ^2 indicate the lack of random effects (a systematic bias), high values indicate systematic differences between the workers.

Hawkins (1991, pp 5, 160) defines the homogeneous exposure group correctly as "a group of workers with identical *probabilities* of exposure to a single environmental agent," (italics added) and as "a group of employees who experience agent exposures similar enough that monitoring agent exposures of any worker in the group provides data useful for predicting exposures of the remaining workers," respectively. This does not imply that the sample mean and variance of a limited series of measurements taken from individuals within the HEG should be equal, nor that mean and variance may not differ significantly between the extreme series.

I encourage Rappaport and Kromhout to reanalyze their unique database, using a criterion that takes into account the effects of limited serial sampling within an HEG. Also I would advise them to compare the effectiveness of such a criterion with the stipulated methods that already exist in literature. Let us keep the statistical analysis of industrial hygiene data as simple as possible and prevent introducing new methods that are at most as effective as existing methods.

I hope the "publication in preparation" will provide insight on the relation between ${}_{B}R_{.95} = 2$ and the random effects model and will explain why about three-quarters of the HEGs with a ${}_{B}R_{.95} < 2$, have a negative "between worker" variance ($\sigma^2 < 0$!).

Industrial hygienists are advised to test inhomogeneity in measurement series among homogeneous exposure groups using classical methods such as control charts, lognormal probability plots, and omnibus tests for normality (see Shapiro 1990, Hawkins 1991).

Theo M.L. Scheffers
DSM Chemical Company

BIBLIOGRAPHY

Abramowitz, M. and I.A. Stegun: *Handbook of Mathematical Functions.* 5th ed. New York: Dover Publications, Inc., 1970.

Guest, G.I., J.W. Cherrie, R.J. Gardner, and C.D. Money: *Sampling Strategies for Airborne Contaminants in the Workplace.* [Technical Guide No. 11] British Occupational Hygiene Society, 1993.

Hawkins, N.C., S.K. Norwood, J.C. Rock (eds.): *A Strategy for Occupational Exposure Assessment.* Akron, Ohio: American Industrial Hygiene Association, 1991.

Health & Safety Executive: Monitoring strategies for toxic substances. [Guidance note EH 42] *Environ. Hyg. 42*:1–12 (1989).

Kromhout, H., E. Symanski, and A.M. Rappaport: A comprehensive evaluation of within- and between-worker components of occupational exposure to chemical agents. *Ann. Occup. Hyg. 17*:253–270 (1993).

Land, C.E.: Hypothesis tests and interval estimation. In *Statistics: Textbooks and Monographs, Vol. 88, Lognormal Distributions.* Ed. by Edwin L. Crow and Kunio Shimizu. New York: Marcel Dekker, 1988 pp. 87–112.

Rappaport, S.M.: Assessment of long-term exposures to toxic substances in air. *Ann. Occup. Hyg. 15*:61–121 (1991).

Rappaport, S.M., H. Kromhout, and E. Symanski: Variation of exposure between workers in homogeneous exposure groups. *Am. Ind. Hyg. Assoc. J. 54*:654–662 (1993).

Shapiro, S.S., M.B. Wilk: An analysis of variance test for normality. *Biometrika 52*: 591–611 (1965).

Shapiro, S.S.: How to test normality and other distributional assumptions. In *The ASQC Basic References in Quality Control: Statistical Techniques*, 2nd ed., vol. 3. Ed. by S.S. Shapiro and E.F. Mykytka. The American Society for Quality Control, 1990.

Dear Sir:

The article on "Variation of Exposure Between Workers in Homogenous Exposure Groups" by Rappaport et al. [AIHAJ, 54:654–662 (1993)] is profoundly thought provoking and a significant contribution to the literature. However, I am concerned that the reader may take away the mistaken impression that establishing exposure groupings by the observational approach is fundamentally flawed and of no utility.

The sampling approach utilized by Rappaport et al. (random selection of workers followed by random selection of monitoring periods) is well suited for establishing exposure groupings in epidemiological studies, but it is not well suited for identifying specific individuals who are the most highly exposed. While the sampling approach is inherently more accurate than the observational approach, it requires a large number of samples to be collected, and in most organizations there are limited resources available for extensive sampling campaigns. The AIHA Exposure Assessment Strategies Committee (EASC) position is that an industrial hygienist's time is well spent in performing an initial comprehensive assessment of the workplace, the workforce, and the environmental agents in an effort to establish and prioritize exposure groups for monitoring.

One of the advantages of the observational approach is that conditions leading to high exposures can be identified and resolved qualitatively without proceeding to a quantitative assessment step. In addition, much of the industrial hygiene monitoring done today specifically targets the most highly exposed workers to evaluate compliance with OSHA PELs or other occupational limits (OELs). The sampling approach proposed by Rappaport et al. acknowledges a role for the observational approach but argues a comprehensive observational step is more appropriate for follow-up investigations after a first round of monitoring indicates there is a problem.

The EASC advocates the industrial hygienist specify the sampling strategy (i.e., representative, worst case) for the specific purpose (i.e., baseline, compliance, diagnostic). Exposure groupings by process/job/agent and perhaps task are determined initially by qualitative assessment methods utilizing professional judgment (i.e., identification of process emission points, handling practices, agent toxicity, etc.) followed by random selection of workers for sampling within the exposure grouping. Fundamental to all exposure monitoring is an initial walk-through survey to assist in the gathering of this qualitative information. It is further promoted that periodic reevaluation of the exposure groupings be performed to confirm their stability and to refine their make-up.

Rappaport et al. raises two valuable concerns that all industrial hygienists need to consider. First, if all workers in a group are assumed to have the same exposure distribution (a homogenous distribution) then the computed average value is treated as every worker's mean exposure level. As noted by Rappaport et al., this aggregate data analysis ignores the true differences in arithmetic mean exposure levels that exist between workers in the group and consequently the industrial hygienists may fail to recognize those workers having unacceptably high exposures. It is further worth noting that some companies set their internal action limits at some fraction of the occupational exposure limit in order to address the issue of outlying exposures at the tail of the distribution. Interestingly, Rappaport et al. raises at least one "homogenous exposure group" with a range of exposures among workers greater than 100-fold. Clearly there is a need to continually refine exposure groups as information becomes available.

Second, the presence of substantial between-worker variability in a group that basically performs the same tasks suggests that individual work practices may be important, in which case exposure reduction could be achieved by developing and implementing standard work practices that minimize exposure while performing the task. Alternately, Nicas and Spear[1] suggest that among workers who perform the same tasks, differences in the overall fraction of time spent performing specific tasks could be a major cause of between-worker exposure variability. In either case it is likely that only through extensive sampling is a problem going to be detected, and analysis of this variability can only be done if there are repeated samples taken on the same workers. This is an expensive proposition that the industrial hygiene profession has yet to address.

EASC largely accepts Professor Rappaport's critique of the term "homogenous exposure group" (HEG) as it is statistically defined. The committee's working definition of an HEG is "a group of employees who experience agent exposures similar enough that monitoring agent exposures of any worker in the group provides data useful for predicting exposures of the remaining workers. Such groups are used in stratified sampling of workplace exposures, thereby improving the power of decision tools.''[2] EASC is currently working on a second edition of the monograph *A Strategy for Occupational Exposure Assessment*, and is expanding the discussion of exposure groupings and the observational and sampling approaches.[2]

Despite the difference in emphasis between Rappaport et al. and EASC, we both agree that formation of exposure groupings is essential, that the observational and sampling approaches each have their advantages and disadvantages, and that continual refinement of exposure groupings is important. It should be apparent from these comments that we see exposure assessment methodologies as an evolving area of industrial hygiene practice. The EASC seeks and supports an open debate on these important issues. In this regard we recommend to the reader the recent editorial in the *Annals of Occupational Hygiene* titled ''Shifting Concepts in Assessment of Occupational Exposures.''[3] Professor Rappaport's insights have much to offer the industrial hygiene profession, and we invite him to participate in integrating his ideas and understanding into our effort.

Christopher J. Cole, CIH, CSP
Chair, EASC
Bionetics Corp.

REFERENCES

1. **Nicas, M. and R.C. Spear:** Task based statistical model of a worker's exposure distribution. *Am. Ind. Hyg. Assoc. J. 54*: 211–227 (1993).
2. **Hawkins, N.C., S.K. Norwood, J.C. Rock, eds.:** *A Strategy for Occupational Exposure Assessment.* Akron, Ohio: American Industrial Hygiene Association, 1991.
3. **Burdorf, A.:** Shifting concepts in assessment of occupational exposures. *Ann. Occup. Hyg. 37*:447–450 (1993).

Dear Sir:

The article by Rappaport et al. [AIHAJ 54: 654–662 (1993)] makes a very important contribution to a growing and fruitful debate concerning the best way to group workers (and exposure measurements) for analysis.

Without detriment to this important article, however, I have three comments that will hopefully contribute to this debate:

(1) The concept of ''homogeneity'' may well require *different definitions* depending on the purpose and the particular type of grouping of exposures involved. For example, the type and degree of homogeneity that might be desirable for grouping workers in an epidemiologic study seeking to elucidate the effects of low-level exposures to benzene or formaldehyde would necessarily be more elegant than the homogeneity required by a regulator analyst seeking to group workers for a population risk assessment in a particular industry. These two needs for grouping of exposure data have different purposes, they involve different levels of aggregation of data, and they have different needs for statistical homogeneity (however defined).

(2) The article presents convincing evidence that ''some workers [in traditionally defined homogenous exposure groups] are consistently exposed to much greater concentrations than their coworkers,'' possibly because of different work tasks and practices. The authors' assertion, however, that such differences in tasks or work practices ''do not lend themselves to engineering solutions'' does not seem justified. It is both reasonable and common that exposures for different tasks and practices, when they have their origin in the same source, would be quite amenable to engineering (source) controls, or even to control of the pathway between the source and the different workers. Current methods of grouping and sampling can address this problem, despite the issues of variability.

A recent article by K.S. Thind et al.[1] describing pentachlorophenol exposures to electric utility linemen is a good albeit somewhat simplistic example. Their findings clearly show what the workers themselves had described: linemen with less seniority were required to do more pole climbing and therefore had higher exposures. In this instance, the source of exposure was the chemical in the poles, the route was skin absorption, and the exposures of the entire group would be reduced by increased use of protective equipment. Yet the exposure measurements would probably be consistent with the findings of Rappaport et al.

(3) When referring to ''investigating job-specific tasks and practices'' as a means to develop improved observational schemes for grouping, the authors pessimistically dismiss the approach ''because there is simply no convenient point where one can stop observing an increasingly subtle array of tasks.'' I prefer to see the glass half-full. The information about exposure determinants that is currently collected with most exposure measurements is woefully inadequate, and a strong argument could be made that better collection of such information would allow far better grouping, analysis, and utilization of exposure data. It would also allow better modeling and many other improved uses of exposure data. One step in that direction would be to agree on the determinants we need to collect, how we will define them, and how we will code them. It is worth a try.

Manuel R. Gómez, MS, CIH
United States Environmental Protection Agency

Dear Sir:

We read with interest the thoughtful letters from Cole, Gomez, and Scheffers regarding the article, ''Variation of Exposure Between Workers in Homogeneous Exposure Groups.''[1] Since we had hoped to provoke a debate on the concept of homogeneous-exposure groups or HEGs, it was gratifying to see such interest in our paper.

We agree with Cole that readers should not conclude from our work that the observational approach is fundamentally flawed. Indeed, some type of observational grouping is essential to any plan for assessing exposures. We expect that industrial hygienists would use major factors in classifying workers, primarily those based on job, location, tasks, and types of controls. However, the assumption that exposure would be ''homogeneous'' in all such groups is questionable in light of our findings and should be investigated. While Cole agrees that statistical sampling provides a more accurate basis for grouping workers, he is concerned that too many measurements would be needed to make this a viable option. Rather, he encourages industrial hygienists to perform ''an initial comprehensive assessment of the workplace'' so that ''conditions leading to high exposures can be identified and resolved qualitatively.'' We remain skeptical of this approach in general because of the enormous variation in exposure associated with industrial processes, except possibly for continuous-indoor operations. For example, in our companion article[2] we reported that exposure typically varied by about 100-fold within persons in a given HEG when the industrial process was either 'outdoors' (${}_w\hat{R}_{.95} = 104$) or 'intermittent' (${}_w\hat{R}_{.95} = 94$). (Note that ${}_w\hat{R}_{.95} = e^{(3.92\hat{\sigma}_w)}$ represents the estimated ratio of the 97.5th to the 2.5th percentile of eight-hour

TWAs for any given worker in the group). Since this variation confounds the ability to discern differences between persons, it seems rather fanciful to expect such situations to be "resolved qualitatively" unless something can be done to reduce the day-to-day variability associated with σ_w^2. Therefore, we encourage hygienists to devote more resources to measuring exposures per se rather than to "initial comprehensive assessment(s)."

We concur with Gomez that the definition of "homogeneity" should suit the application. We also agree that engineering controls can reduce even those exposures associated with tasks or work practices (enclosure of a source of contamination comes to mind) but don't believe that such interventions would necessarily prove effective given the propensity of human beings to find novel ways of doing the same job. Regarding his last point (which is similar to Cole's argument about initial comprehensive assessments), however, we remain unconvinced that the hygienist is better off by collecting more and more information about tasks, activities and the like, prior to making personal measurements, because some of the data are likely to be extraneous to the determinants of exposure. Since the relevance of such information can only be evaluated in the context of measurements, generally it would be desirable to identify major tasks or activities as exposures are being measured. A good illustration of this approach has recently been reported by Kromhout et al.[3]

Scheffers' criticisms focus on the influence of random error on the identification of "homogeneous" groups. First, he takes exception to our definition of a "uniformly exposed group" in which the exposures of 95% of the persons lie within a two-fold range (i.e., where ${}_BR_{95} \le 2$).[4] (Note that ${}_BR_{95} = e^{(3.92\sigma_B)}$ represents the ratio of the 97.5th to the 2.5th percentile of the workers in the group). Although this provides a useful benchmark for gauging whether an observational group is "homogeneous," it is not sacred, and the debate should not be restricted to establishing a particular value. In fact, we reported point estimates of ${}_BR_{95}$ (designated ${}_B\hat{R}_{95} = e^{(3.92\hat{\sigma}_B)}$) that varied between 1 and 2000 for 183 HEGs (shown as Figure 2 in our paper[1]). Surely, most hygienists would agree that somewhere in this vast range, the HEG is not "homogeneous!"

Sheffers goes on to suggest that our analyses were flawed because we failed to account for the random error in the point estimates of ${}_BR_{95}$ that we reported.[1,2] In other words, he implies that our HEGs might have been homogeneous, but we overestimated the true ${}_BR_{95}$s, thereby making the situation appear more heterogeneous than it really was. He bolsters his case by investigating some data generated by Monte Carlo techniques. Since Scheffers kindly supplied us with a file containing one of his data sets (for k = 5 workers per group and n = 2 measurements per worker), we were able to analyze his artificial HEG as well as our real ones to determine the impact of random error on estimation of ${}_BR_{95}$.

RANDOM ERROR IN ESTIMATING THE BETWEEN-PERSON VARIANCE

In the paper[1] we reported 183 values of the estimated between-person component of variance ($\hat{\sigma}_B^2$) and the corresponding values of ${}_B\hat{R}_{95}$. Values of $\hat{\sigma}_B^2$ were obtained using the one-way random-effects ANOVA model with unbalanced data according to standard methods.[5] As discussed in the paper,[1] the model was fit to the natural logarithms of the air concentrations.

Scheffers' Simulated Data

Scheffers presented two sets of simulated data for the situations where k = 5, n = 2 and where k = 10, n = 5. Each set consisted of 1000 samples that had been generated from a lognormal distribution, for which $\mu_y = 0$, $\sigma_y^2 = \sigma_w^2 = 1.0$, and $\sigma_B^2 = 0$. Since $\sigma_B^2 = 0$ (${}_BR_{95} = 1$), the assumed groups were truly "homogeneous." Yet, any single sample of data from the distribution could yield a point estimate ($\hat{\sigma}_B^2$) that is positive, or equivalently an ${}_B\hat{R}_{95} > 1$, because of random error. Thus, Scheffers makes a good point, which we will come back to shortly. Meanwhile we will address some inconsistencies in his analyses.

Although it is reasonable to investigate the effects of random error on estimates of ${}_BGSD = e^{\sigma B}$ by analyzing simulated data, Scheffers' use of nonstandard methods led to substantial bias in estimating both ${}_BGSD$ and ${}_wGSD = e^{\sigma W}$. This is obvious from a statistical perspective, since proper methods applied to the data, simulated under the conditions where $\sigma_B^2 = 0$ and $\sigma_w^2 = 1$, would yield average values of 1 and 2.718 for estimates of ${}_BGSD$ and ${}_wGSD$, respectively, when large numbers of samples are analyzed, as is the case here. Since Scheffers' estimates of ${}_BGSD$ and ${}_wGSD$ differed substantially from these average values, there is a flaw in his method of estimation. Although we have some theories as to why Scheffers did not reach the correct outcomes, the fact that he did not indicates that his specific conclusions cannot be relied on. (Note: by applying standard methods we obtained average values of $\hat{\sigma}_B^2$ near 0 and $\hat{\sigma}_w^2$ near 1 from one of Scheffers' files; thus, it appears that his methods were biased rather than his simulated data). Although it would be possible to question the properties of the estimator ${}_B\hat{R}_{95}$ used in the paper,[1] this estimator is reasonable and not subject to the severe bias to which Scheffers' estimator falls victim.

Despite the failings of Scheffers' methods, he correctly identifies random error as a potential problem in estimating ${}_BR_{95}$. Each ANOVA estimate of σ_B^2 is subject to imprecision, the amount of which increases when k and n get small and σ_w^2 becomes large. This variability is magnified when transferred to the derived estimate of ${}_BR_{95}$. To illustrate this, we computed 1000 ANOVA estimates for σ_B^2 from Scheffers' data set in which k = 5, n = 2. Even though 55% of the 1000 estimates of σ_B^2 were less than or equal to zero, 42% were greater than 0.031 (${}_B\hat{R}_{95} = 2.0$), 18% were greater than 0.587 (${}_B\hat{R}_{95} = 10$), and 2% were even greater than 0.998 (${}_B\hat{R}_{95} = 50$). Thus, although one would reach the "correct" conclusion, i.e., that the group was "homogeneous," based on more than half of the samples, one might also suspect significant heterogeneity on the basis of about 20% of the samples. Of course, by collecting larger samples, such errors would be reduced, as is always the case with statistical analyses. In fact, Scheffers' data set depicts the worst case since k = 5, n = 2 represents the minimum sizes used in our analyses,[1] where the median numbers of measurements and workers per group were 28 (with a range of 11–5076) and 8 (with a range of 5–62), respectively.

Confidence Intervals for $\hat{\sigma}_B^2$

In order to determine the potential impact of random error on the results presented in our paper,[1] we estimated 95% confidence intervals for σ_B^2 in our 183 data sets, using an approximation provided by Burdick and Eickman[6] for unbalanced data. We then computed the corresponding intervals for ${}_BR_{95}$, and

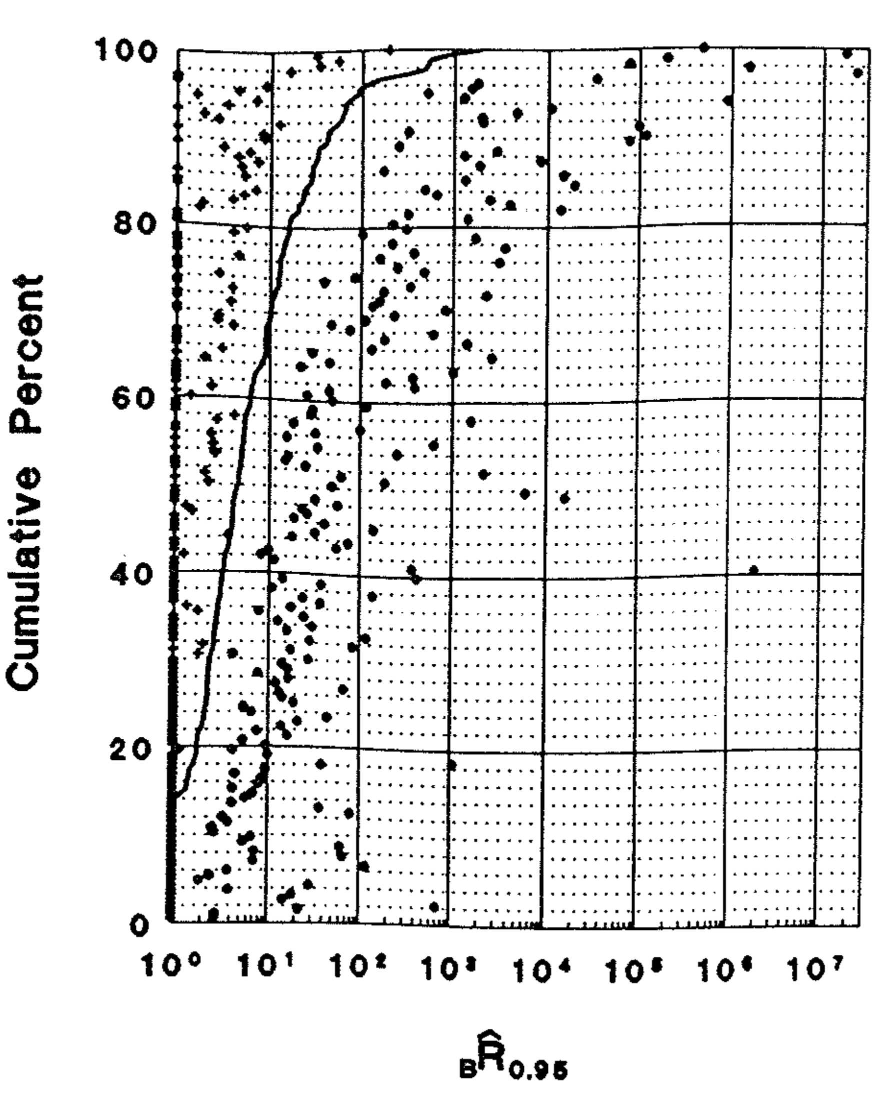

FIGURE 1. The solid curve depicts the empirical cumulative distribution function of $_B\hat{R}_{95}$ for 183 HEGs.[1] Individual points represent lower (+) and upper (•) bounds on a 95% confidence interval for $_BR_{95}$ for each group.

plotted these upper and lower limits along with the 183 original values of $_B\hat{R}_{95}$ in Figure 1. Note that if $\hat{\sigma}^2_B$ were less than zero for a given group, we set $\hat{\sigma}^2_B$ (and the lower confidence limit) to zero, following a defensible statistical practice.[5]

The figure, which covers more than eight orders of magnitude, clearly shows that estimates of $_BR_{95}$ can be extremely imprecise, as suggested by Scheffers. However, the limits also allow us to make inferences about the between-person variation in HEGs. Focusing first on the lower limits, we can say with confidence that 30% of the groups were not "homogeneous," since these lower limits correspond to values of $_BR_{95} > 1$. Likewise, by referring to the upper confidence limits we cannot rule out the possibility that 52% of the HEGs have $_BR_{95} > 50$, suggesting extreme heterogeneity.

CONCLUSION

Figure 1 helps to crystallize the debate about HEGs. If one adheres to the conventional wisdom that observational groups are always "homogeneous," it should be disquieting to observe a statistically significant lack of homogeneity in 30% of our HEGs. On the other hand, those who are skeptical about observational groups in general can note that more than half of our HEGs could contain workers with more than a 50-fold range of exposures. Either way, it seems prudent for industrial hygienists to adopt sampling practices that allow them to estimate the components of variance, so that they can gauge for themselves whether or not a particular group is "homogeneous."

S.M. Rappaport, Ph.D.
E. Symanski
R. Lyles
University of North Carolina

H. Kromhout, Ph.D.
Wageningen Agricultural University

REFERENCES

1. **Rappaport, S.M., H. Kromhout, and E. Symanski:** Variation of exposure between workers in homogeneous exposure groups. *Am. Ind. Hyg. Assoc. J. 54*:654–662 (1993).
2. **Kromhout, H., E. Symanski, and S.M. Rappaport:** A comprehensive evaluation of within- and between-worker components of occupational exposure to chemical agents. *Ann. Occup. Hyg. 37*:253–270 (1993).
3. **Kromhout, H., P. Swuste, and J.S.M. Boleij:** Empirical modeling of chemical exposure in the rubber-manufacturing industry. *Ann. Occup. Hyg. 38*:3–22 (1994).
4. **Rappaport, S.M.:** Assessment of long-term exposures to toxic substances in air. *Ann. Occup. Hyg. 35*:61–121 (1991).
5. **Searle, S.R., G. Casella, and C.E. McCulloch:** *Variance Components.* New York: John Wiley and Sons, 1992.
6. **Burdick, R.K. and J. Eickman:** Confidence intervals on the among group variance component in the unbalanced one-fold nested design. *J. Stat. Comp. & Simul. 26*: 205–219 (1986).